

Władysław Marek Kolasa

Formaty dokumentów w bibliotekach cyfrowych

1. Wprowadzenie

Format dokumentu w potocznym rozumieniu zazwyczaj kojarzy się z rozszerzeniem pliku, który wskazuje na aplikację potrzebną do jego uruchomienia. Takie pojmowanie jest oczywiście słuszne, lecz niepełne, gdyż w ścisłym rozumieniu format obejmuje całą gamę abstrakcji związanych z pojęciem dokumentu cyfrowego. Do precyzyjnego określenia jego zakresu niezbędne jest zatem zdefiniowanie dokumentu elektronicznego. W literaturze istnieje wiele sposobów jego definiowania, np. informatologiczne¹, prawne², czy administracyjne³, jednak z technicznego punktu widzenia odpowiednie wydaje się określenie zapisane w pierwszym arkuszu normy *Informacja i dokumentacja - zarządzanie dokumentami* (PN-ISO 15489-1:2006). Dokument elektroniczny jest tam zdefiniowany jako: *dokument istniejący w postaci elektronicznej, dostępny za pośrednictwem techniki komputerowej*. Definicja akcentuje zatem jego dwie dystynktywne cechy: trwałość i możliwość lektury. Ważne jest więc, aby dane były zapisane i przechowywane na elektronicznym nośniku i aby można je było odtworzyć w formie umożliwiającej ich lekturę. Nie ma przy tym znaczenia sam nośnik informacji (np. pendrive, dysk twardy, serwer, sieć) ani rodzaj zawartych danych (tekst, grafika, multimedia), istotny jest jedynie fakt, aby dane te mogły zostać wielokrotnie odtworzone i aby mógł zachodzić akt percepcji. Z praktycznego punktu widzenia warto poczynić jednak drobną dygresję: w niniejszym opracowaniu będą nas interesowały wyłącznie dokumenty w formie skończonej (zamkniętej), przeznaczone do publicznego rozpowszechniania – czyli publikacje elektroniczne. Takie ujęcie wychodzi naprzeciw bibliotecznemu pojmowaniu zbiorów i pozwoli uniknąć problemów, jakie rodzą nowe typy dokumentów (np. blogi, posty).

¹ Dokument elektroniczny (komputerowy, cyfrowy) - dokument z sygnałem utrwalonym na nośniku elektronicznym. Mogą to być komunikaty różnych języków naturalnych w subkodach akustycznych i graficznych, inne komunikaty graficzne jak obrazy, wykresy, animacje, filmy, komunikaty dźwiękowe - *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*. Warszawa 2002, s. 50.

² Dokument elektroniczny - każdy przedmiot lub zapis na komputerowym nośniku informacji, z którym jest związane określone prawo albo który ze względu na zawartą w nim treść stanowi dowód prawa, stosunku prawnego lub okoliczności mającej znaczenie prawne - *Kodeks Karny*, Dz.U. nr 140/1997, poz. 553.

³ Dokument elektroniczny - zbiór danych wprowadzonych lub przechowywanych na dowolnym nośniku przez system informatyczny lub podobny układ, które mogą być odczytane lub wyświetlone przez osobę lub przez tego rodzaju system lub układ, a także wszelkiego rodzaju prezentacja i wszelkiego rodzaju przedstawienie tych danych w formie drukowanej lub innej - *Decyzja Komisji Europejskiej* nr 2004/563/WE.

Przy bliższej analizie rozważanego pojęcia można wskazać na kilka jego cech szczegółowych. Każdy dokument elektroniczny składa się z czterech abstrakcji: pierwszą jest zawartość, czyli jego warstwa znaczeniowa (semantyczna); drugą – struktura logiczna, odpowiadająca z jakich jednostek dokument jest złożony (oraz z jakimi powiązany); trzecią – struktura fizyczna (czyli **format danych**) oraz czwartą – tzw. kontekst, rozumiany jako metadane, określające w jakiej postaci dokument będzie prezentowany. Kluczowe znaczenie dla dalszych rozważań ma struktura fizyczna, którą można określić, jako wynik przetworzenia informacji zawartych w dokumencie elektronicznym na dane w układzie binarnym. Zasadniczym polem dalszych rozważań będą więc problemy kodowania i innych transformacji jakim podlega dokument oraz stosowane w tym zakresie standardy. Warto dodać, że operacje te są odbywają się w środowiskach języków programowania wysokiego poziomu, zazwyczaj w aplikacjach macierzystych (np. Adobe Acrobat). Z informatycznego punktu widzenia zagadnienia formatu danych można jeszcze podzielić się na dwie grupy: pierwsza obejmuje atrybuty fizyczne (np. algorytmy kompresji, kodowane), druga – atrybuty typograficzne (np. formatowanie, czcionka).

Jakkolwiek na użytek praktyki stworzono liczne typologie formatów danych, jednak brakuje schematu uniwersalnego. W przypadku bibliotek cyfrowych rolą taką może jednak pełnić wykaz typów MIME⁴, specyfikacja zarządzana przez Internet Assigned Numbers Authority (IANA) i wspierana przez Konsorcjum W3C. Standard ów odnosi się wprawdzie do protokołów komunikacyjnych, lecz równolegle precyzyjnie definiuje także związane z nimi formaty i wskazuje na oficjalne specyfikacje. Wśród dziewięciu głównych grup wykaz MIME wylicza m.in. formaty: dźwiękowe (audio), graficzne (image), tekstowe (text), wideo (video) oraz związane z programami (application). Z punktu widzenia repozytoriów cyfrowych warto zwrócić uwagę na kilka najważniejszych; w grupie tekstowej będą to m.in.: czysty tekst (plain text), HTML i XML; w grupie audio: MPEG (np. MP3), Windows Media Audio, RealAudio i WAV; w grupie graficznych: GIF, JPG, PNG, TIFF, DJVu; w grupie wideo: MPG-1, MP4, QuickTime, Windows Media Video; spośród innych (application): JavaScript, Adobe Flash, strumień (octet-stream) oraz PDF. Warto podkreślić, że wskazane specyfikacje odwołują się także do licznych standardów ogólnych (np. kodowania znaków, barwy i in.), które są odrębnie definiowane i opisane. Szczegółowych informacji na ich temat należy poszukiwać w odpowiednich źródłach (np. International Organization for Standardization, The Unicode

⁴ MIME Media Types: <http://www.iana.org/assignments/media-types/index.html> [2011.04.10]; Internet Media Type registration, consistency of use TAG Finding 3 June 2002 (Revised 4 September 2002) - <http://www.w3.org/2001/tag/2002/0129-mime> [2011.04.10].

Consortium, International Color Consortium i in.). Istotną rolę w poprawnym przetwarzaniu i interpretacji standardów przez systemy teleinformatyczne pełnią nadto osadzone w dokumentach metadane techniczne, które również mogą być odrębnie zdefiniowane, np. standardy zapisu informacji w plikach graficznych IPTC (International Press Telecommunications Council)⁵ czy EXIF (Exchangeable Image File Format)⁶. Trendem ostatnich lat jest natomiast otwarte podejście do tej kwestii, czyli definiowanie kontenera metadanych w formie XML, co sprawia, że może być on dołączony do opisywanego obiektu na wiele sposobów i zawierać różne elementy; takie rozwiązanie przyjęto np. w specyfikacji MPEG-7, dedykowanej dla opisu treści multimedialnej⁷.

2. Formaty w cyklu tworzenia publikacji cyfrowej

W praktyce bibliotek cyfrowych wybór formatu zależy od kilku kwestii, wśród których czołową rolę odgrywa stopień przetworzenia i przeznaczenie materiału. Zazwyczaj wyróżnia się trzy etapy cyklu przetworzenia i odpowiadające im formaty. Na etapie pierwszym, który obejmuje: przygotowanie, digitalizację, korekty, montaż, obróbkę i porządkowanie materiału podlega wielu przekształceniom i jest zachowywany w rozmaitych formatach **pośrednich** (przeważnie natywnych); istotną rolę odgrywa tu nie tyle sam format ile poprawny system wersjonowania plików i ich notacji (nazewnictwa). W etapie drugim finalny wersja przetworzonych i zmontowanych plików staje się podstawą wersji archiwalnej (master), którą należy przygotować w odpowiednim formacie **archiwalnym**. Podstawową funkcją tej kopii jest długotrwałe przechowywanie publikacji z zachowaniem możliwie najwyższych parametrów. Na etapie trzecim w oparciu o kopię master tworzy się publikację w wersji przeznaczonej do prezentacji w sieci, czyli publikację w formacie **prezencyjnym**. Ostatnia wersja jest zazwyczaj owocem kompromisu pomiędzy trzema parametrami: niewielkimi rozmiarami, akceptowalną jakością i wysokim poziomem zorganizowania. Aby osiągnąć ów kompromis należy skorzystać z któregoś z formatów oferujących skuteczne algorytmy kompresji stratnej (np. JPG dla grafiki rastrowej). Należy podkreślić, że dobór formatu prezencyjnego jest zadaniem dość trudnym, gdyż uzyskane efekty silnie zależą od rodzaju publikacji, w szczególności takich jej parametrów jak: ilość, głębokość i rozmieszczenie kolorów, rozmiary fizyczne, rodzaj prezentowanej treści, a czasem nawet użytych znaków (np. symbole matematyczne lub wzory chemiczne).

⁵ IPTC: <http://www.iptc.org/> [2011.04.10].

⁶ EXIF.org: <http://exif.org/specifications.html> [2011.04.10].

⁷ MPEG-7 Overview: <http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm> [2011.04.10].

3. Formaty archiwalne

Formaty archiwalne (FA) w porównaniu do formatów prezencyjnych nie są ani zbyt skompilowane, ani nie wymagają większej wiedzy na etapie użytkowania. Z drugiej strony jedynie kilka standardów aspiruje do pełnienia tej roli (głównie PDF/A i TIFF). Jeśli – przypomnijmy – rolą kopii master jest długotrwałe przechowywanie publikacji z zachowaniem możliwie najwyższych parametrów, stąd łatwo o wniosek, że jedną ważną cechą jest wierne odwzorowanie oryginału. Aby osiągnąć ten cel deponuje się zwykle pliki archiwalne z zastosowaniem maksymalnych atrybutów, częstokroć stosując wielkości nadmiarowe. Plastycznie ilustruje to przykład grafiki rastrowej: jakkolwiek maksymalna głębia dla w modelu RGB wynosi 24 bity (true color), skany do celów archiwalnych wykonuje się w głębi 36 lub 48-bitowej, stosując jednocześnie wyższą rozdzielczość (np. nie 300, lecz 600 lpi [line per inch]); identyczne podejście przyjmuje się zresztą do dźwięku i obrazy (próbki nadmiarowe) oraz innych rodzajów danych. Efektem zastosowania wspomnianych danych nadmiarowych są gigantyczne rozmiary plików master, gdyż ich objętość rośnie bądź proporcjonalnie do zmiany parametrów (np. głębi lub próbkowania), bądź proporcjonalnie do ich kwadratu, np. gdy zwiększamy rozdzielczość. Problem objętości nabiera szczególnego znaczenia, jeśli uwzględnić fakt, że zachowanie wysokiej jakości pociąga za sobą konieczność użycia plików bez kompresji lub stosowania algorytmów kompresji bezstratnej. W efekcie kopie archiwalne mają bardzo duże objętości i do przechowywania wymagają specjalnych urządzeń (zwykle macierzy dyskowych). Wymagania stawiane kopiom archiwalnym szczegółowo opisują różne dokumenty, np. *Decyzja Nr 25 Naczelnego Dyrektora Archiwów Państwowych*⁸ lub odpowiednie fragmenty dokumentacji PREMIS, rozwijanej przez OCLC i Bibliotekę Kongresu USA⁹.

Wśród kilku formatów stosowanych do zachowywania macierzy archiwalnych główną rolę dla grafiki rastrowej pełni specyfikacja **TIF** (Tagged Image File Format). O jej przydatności zdecydowały następujące cechy formatu: obsługa wszystkich rodzajów głębi i przestrzeni kolorów (RGB, CMYK), możliwość bezstratnej kompresji (algorytm LZW), możliwość przechowywania ścieżek i kanałów alfa oraz profili koloru i komentarzy tekstowych, a nade wszystko to, że jest ona standardem międzyplatformowym¹⁰. Jedynym

⁸ Decyzja Nr 25 Naczelnego Dyrektora Archiwów Państwowych z dnia 30.12.2005 r. w sprawie standardów sporządzania cyfrowych kopii materiałów archiwalnych w archiwach państwowych - http://www.archiwa.gov.pl/images/stories/file/pdf/dec_25_2005.pdf [2011.04.14].

⁹ PREMIS maintenance activity: <http://www.loc.gov/standards/premis/> [2011.04.14].

¹⁰ Specyfikacja TIF rev. 6.0: <http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf> [2011.04.14]

problemem, jaki pojawia się przy okazji stosowania TIF, jest pytanie czy stosować pliki wielostronicowe? Praktyka uczy jednak, aby nie korzystać z tej możliwości, gdyż na dłuższą metę może to powodować błędy w przetwarzaniu, o problemach z ekstrakcją pojedynczych stron nie wspominając. Jedynym mankamentem stosowania TIF jest zatem niski poziom zorganizowania macierzy (każdy plik osobno). Aby zapobiec możliwym problemom należy przyłożyć wiele uwagi do właściwej notacji i systemu nazewnictwa, gdyż muszą one być poprawnie obsługiwane w różnych systemach (Windows, Unix, MacOS i in.). Należy więc zadbać, aby każda jednostka była zdeponowana w odrębnym folderze i aby mechaniczne sortowanie plików było zgodne z naturalną kolejnością stron finalnej publikacji. Druga – i bodaj ważniejsza kwestia – wiąże się z wykazem dozwolonych znaków. Ze względów bezpieczeństwa do nazywania plików należy stosować tylko znaki z bezpiecznej puli. W praktyce są to litery alfabetu łacińskiego (najlepiej małe), cyfry oraz znaki „-” (minus) i „_” (podkreślenie). Poprawne sortowanie z kolei wymaga, aby nazwy plików były równej długości, a indeks cyfrowy był uzupełniony zerami (np.: gazeta_polska_1936_033_00014).

Drugim formatem, który kwalifikuje się do długoterminowego przechowywania danych jest **PDF/A** (Portable Document Format w wersji A), czyli specjalna implementacja PDF uregulowana w 2005 roku normą ISO 32000-1¹¹. Plik PDF pełni w tym zastosowaniu rolę kontenera, tzn. pozwala na zaimportowanie i osadzenie w pliku dowolnych danych (z wyjątkiem audio/wideo). Zasadniczo stosuje się go w dwu przypadkach: do grafiki rastrowej zamiast TIF (wówczas pliki zyskują wyższy poziom organizacji – strony, bookmarki) oraz dla dokumentów testowych (najczęściej biurowych, w tym listów). W ostatnim zastosowaniu PDF/A sprawdza się doskonale, gdyż podczas importu zachowuje wszelkie niezbędne komponenty (w tym np. użyte czcionki i obrazy), gwarantując, że będą one miały identyczny wygląd nawet na innej platformie. Jest to wyjście naprzeciw istotnemu problemowi zgodności wersji dokumentów biurowych, które – jak uczy doświadczenie – z uwagi znaczną labilność wersji w ramach tego samego programu (np. MS Word) i dużej liczbie formatów natywnych (różne edytory) są bardzo trudne do poprawnej archiwizacji. Jedyne problem, jaki dotyczy zastosowania PDF/A wiąże się z odzyskiem dokumentu w trybie do edycji. Operację taką można bezstratnie wykonać dla formatów: TIF, JPG, PNG, TXT, RTF, Word, XML, HTML; w przypadku innych dokumentów będzie zachodzić konieczność konwersji, która może prowadzić do utraty części informacji (np. formatowania).

¹¹ PDF Standards: <http://pdf.editme.com/PDFA> [2011.04.14].

Archiwizacja innych rodzajów danych, w szczególności multimedialnych wymaga odrębnych standardów. W przypadku profesjonalnych filmów właściwym rozwiązaniem jest format **DPX** (Digital Picture Exchange), który zachowuje maksymalne parametry ścieżek, m.in. informacje dla każdego kanału o gęstości optycznej w 10 bitach i pełną informację o użytych barwach; jego wadą są jednak bardzo duże pliki wyjściowe, wymagające dedykowanych nośników. Standard DPX został opracowany przez Society of Motion Picture and Television Engineers (SMPTE) i opublikowany w 2003 roku jako norma ANSI/SMPTE (268M-2003)¹². W archiwach dźwiękowych można natomiast użyć standardu **ALC** (MPEG-4 Audio Lossless Coding), który opisuje wydana w latach 2005-2006 norma ISO/IEC 14496-3. W przypadku z archiwów z multimediami, dla których wymagania DPX nie mogą być spełnione rolę formatu master dla multimedialnych mogą pełnić popularne kontenery wideo lub audio, np. **AVI** (Audio Video Interleave) lub **WAV** (wave form audio format) z użyciem kodeków bezstratnych (np. FLAC - Free Lossless Audio Codec) lub bez kompresji.

4. Formaty prezencyjne

Formaty prezencyjne (FP) są grupą największą i jednocześnie najbardziej znaną, gdyż to z nimi obcuje bezpośrednio użytkownicy serwisów i repozytoriów. Zasadniczym celem FP jest prezentacja dokumentów w sieci Web, z czego wynikają dalsze konsekwencje. A zatem winny one: po pierwsze – być niezależne od platformy (Windows, Unix, Android, inne); po wtóre – prezentować się identycznie (lub podobnie) na rozmaitych urządzeniach wyjściowych (ekrany, konsole, drukarki, projektory); po trzecie – mieć niewielkie rozmiary; po czwarte – oferować akceptowalną jakość oraz po piąte – cechować się wysokim poziomem zorganizowania (czyli powiązania poszczególnych stron, części). Dwie pierwsze cechy (międzyplatformowość i przenośność) sprawiają, że już na wstępie liczba możliwych FP staje ograniczona do wąskiej grupy specyfikacji zaprojektowanych *ad hoc* do celów wymiany i prezentacji danych (np. TXT, HTML, RTF, JPG, PDF). Ich poprawną obsługę gwarantują poszczególne platformy bądź poprzez dedykowane wersje wtyczek i programów, bądź wspierają wprost na poziomie sterowników i bibliotek. Z drugiej strony postulat ten powoduje, że z grupy FP należy wykluczyć niemal wszystkie formaty dedykowane dla aplikacji, czyli formaty natywne (np. PSD, DOCX). Trudniej natomiast osiągnąć jednocześnie dwa kolejne postulaty: niewielkie rozmiary i wysoką jakość; są to bowiem parametry wzajemnie sprzeczne, gdyż wyższa jakość powoduje zwykle wzrost objętości, i odwrotnie:

¹² Cineon Image File Format Draft: http://www.cineon.com/ff_draft.php [2011.04.14]

niższa – spadek. Ostatecznie więc należy szukać kompromisu pomiędzy niewielkimi rozmiarami a akceptowalną jakością. Konsensus ten umożliwia zastosowanie odpowiednich algorytmów kompresji, które przeważnie mają charakter stratny (np. JPEG, MPEG). Znaczący to, że dokumentu zapisanego z kompresją stratną nie można bez uszczerbku przywrócić do formy pierwotnej. Mimo tej wady kompresja stratna stała się dominującym sposobem rozwiązywania wskazanego dylematu, zaś ciężar odtworzenia formy pierwotnej przeniesiono na format archiwalny. Ostatnim problemem FP jest postulat wysokiego poziomu zorganizowania; w praktyce chodzi o mechanizm nawigacyjny w obrębie dokumentu, jeśli jest on wieloczęściowy lub ma znaczne rozmiary. Problem jest rozwiązywany rozmaicie. W dokumentach tekstowych (TXT, HTML, RTF) lub hybrydowych (PDF) następstwo części jest niejako „zaszyte” w dokumencie lub realizowane metodą hiperłączy i zakładek. W przypadku grafiki rastrowej niezbędny jest mechanizm zewnętrzny, czyli zastosowanie obudowy HTML (rozwiązanie przestarzałe) bądź użycie nawigacji skryptowej (JavaScript), która odwołuje się do nazw plików. Jeszcze inną metodą prezentacji możliwą do zastosowania dla dowolnych danych jest renderowanie dokumentu za pomocą odpowiednich wtyczek (np. Flash); sposób ten z uwagi na duże możliwości i skalowalność stał się szczególnie przydatny w portalach społecznościowych (np. FaceBook).

Biorąc po uwagę wcześniejsze ustalenia i powyższe postulaty można zaproponować następującą typologię formatów prezentacyjnych i wskazać ważniejsze specyfikacje używane przez biblioteki cyfrowe.

Typ dokumentu	Wybrane formaty prezentacyjne i ich stosowanie w bibliotekach cyfrowych	
	Zalecane	Nie zalecane
Tekstowe - nieadnotowane	TXT (UTF-8)	TXT (inne kodowanie)
Tekstowe - adnotowane	HTML, RTF, ODT (UTF-8)	DOC, XLS, inne natywne
Graficzne - rastrowe	JPG, GIF, PNG, JPG2000, TIF (G4)	PSD, inne natywne
Graficzne - wektorowe	Konwersja na rastrowe	CDR, DWG, inne natywne
Hybrydowe	PDF, DJVU	
Audio	MP3, WMA, AAC, MOV, Ogg	Audio CD, WAV
Wideo	MPEG-2, MPEG-4, FLV, MOV, WMV	VCD, DVD, SVCD, DivX, AVI
Animacje	SWF, ActionScript	
Specjalne – mapy	Zoomify	
Specjalne – nuty	Sibelius	
Aplikacyjne	Konwersja na hybrydowe lub audio, wideo	Strumień binarny
Mobilne	LIT, EPUB, PRC, Adobe Digital, PBD	

4.1. Formaty tekstowe

Formaty tekstowe należą do najstarszej formy dokumentów, gdyż istnieją niemal od zarania ery komputerów. Najprostszą ich formą jest tzw. czysty tekst (plain text, **TXT**), czyli plik znakowy nie zawierający formatowania (inaczej: tekst nieadnotowany). Dokumenty tego rodzaju są nieskomplikowane, gdyż składają się wyłącznie ze znaków alfanumerycznych, przestankowych oraz dwu znaków sterujących (koniec wiersza, tabulacja). Ich zaletą jest mała objętość, gdyż liczba bajtów jest w przybliżeniu równa liczbie użytych znaków. Jedynym problemem jaki wiąże się z tymi plikami jest problem kodowania. Pierwotnie stosowano standard ASCII, gdzie znaki kodowano w 7 bitach (0-127), co było wytaczające dla języka angielskiego. Dla innych języków zgodnie z normą ISO/IEC 8859 (16 arkuszy) rozszerzano ASCII do 8 bitów i na dodatkowych pozycjach (128-255) definiowano odpowiednie diakrytyki (język polski był zdefiniowany w arkuszu 2 [tzw. latin 2]). Równolegle funkcjonowały też inne standardy kodowania (np. IBM, Windows). Stan ów był bardzo niewygodny i już w 1991 roku podjęto prace harmonizacyjne. Aktualnie niekwestionowanym standardem obejmującym wszystkie języki świata jest UNICODE (norma ISO/IEC 10646)¹³. Standard pozwala zakodować wszystkie języki świata i wykorzystuje do tego celu zmienną liczbę bajtów (7 do 32). Najczęściej wykorzystywana jest tablica UTF-8; standard ten od połowy pierwszej dekady XXI wieku zagościł w także w systemach operacyjnych i popularnych aplikacjach biurowych wypierając inne specyfikacje. Współcześnie czysty tekst raczej rzadko wykorzystywany jest do prezentacji zasobów w bibliotekach cyfrowych (zwykle jako wariant publikacji); niemniej w starszych serwisach był formą dominującą (np. *Gutenberg Project* czy *Internet Archive*).

Poszerzoną wersją czystego tekstu jest tekst adnotowany (text enriched), czyli wzbogacony o możliwości formatowania lub inne funkcjonalności. Taki zestaw opcji umożliwia w praktyce wszystkie edytory teksów, np. dostępne w popularnych pakietach: MS Office (doc, docx), OpenOffice (sxw), WordPerfect (wpd), Works (doc). Jednak żaden z wymienionych formatów nie aspiruje do miana standardu prezencyjnego, gdyż każdy ma charakter natywny, tzn. w pełni poprawnie funkcjonują jedynie w określonym programie, a nierzadko wersji. Na potrzeby wymiany powstała specyfikacja **RTF** (Rich Text Format) - opracowana w 1987 roku przez Microsoft i sukcesywnie rozwijana; najnowszą wersję standardu (1.9.1) ogłoszono w roku 2008 przy okazji debiutu edytora Word 2007¹⁴. RTF w

¹³ JTC1/SC2/WG2 - ISO/IEC 10646 - UCS - <http://std.dkuug.dk/JTC1/SC2/WG2/> [2011.04.19]

¹⁴ Rich Text Format (RTF) Specification, version 1.9.1 - <http://www.microsoft.com/> [2011.04.19].

odróżnieniu od czystego tekstu definiuje i przechowuje w plikach m.in. informację o użytych czcionkach, ich ułożeniu, atrybutach (np. kolorze), rozłożeniu akapitów, marginesów, tabel oraz parametry osadzonych obiektów (np. rysunków). Użyty tekst może być zapisany w dowolnej stronie kodowej. Mimo wskazanych zalet trudno zaliczyć RTF do formatów dobrze funkcjonujących w internecie. Złożyły się na to cztery powody: pliki RTF mają stosunkowo duże objętości (10-15 razy większe niż TXT); obiekty inne niż tekst (tabele, grafika) są w często zniekształcane podczas odczytu; pliki RTF nie są interpretowane przez przeglądarki oraz – co najważniejsze – mogą zawierać elementy potencjalnie niebezpieczne (makra, wirusy). Drugim formatem o zbliżonej funkcjonalności jest **ODT** (OpenDocument [Text]) – czyli otwarty formatu plików pakietów biurowych¹⁵, rozwijany przez niezależną organizację OASIS (Organization for the Advancement of Structured Information Standards) i uregulowany normą ISO/IEC 26300: 2006. Specyfikacja wyrosła na bazie XML-owego formatu pakietu OpenOffice i oferuje zestaw opcji zbliżony do RTF. ODT jest specyfikacją stosunkowo młodą, lecz bliższa analiza przekonuje, że jest on obciążony licznymi wadami (podobnie jak RTF), to powoduje, że jego zastosowanie w sieci Web jest ograniczone.

Do grupy adnotowanych formatów tekstowych należy także zaliczyć hipertekst, np. **HTML** (HyperText Markup Language)¹⁶ wraz z wariantami (np. arkusze stylów – CSS, XSL) i języki pochodne, szczególnie XHTML (Extensible Hyper Text Markup Language) oraz inne specyfikacje towarzyszące, którymi zarządza World Wide Web Consortium (W3C)¹⁷; nadto wersje specjalne, np. dedykowane dla urządzeń mobilnych, np. WML (Wireless Markup Language) nadzorowane przez Open Mobile Alliance (OMA)¹⁸. Wspólnym rysem hipertekstu jest poziom uboższy względem RTF poziom formatowania (np. brak justowania, dzielenia wyrazów, zaawansowanych tabel, zagnieżdżeń), a jednocześnie lepsza obsługa elementów osadzonych, np. grafiki. Hipertekst nie jest jednak obciążony wadami RTF: pliki są zawsze małe, mniej podatne na przenoszenie wirusów i – co najważniejsze – obsługiwane wprost przez przeglądarki. Jedynym mankamentem jest jedynie labilność wyglądu strony, gdyż przeglądarki w zależności od lokalnych ustawień, wielkości ekranu i użytych komponentów (np. czcionki) inaczej wyświetlają zawartość. Czasami jest to cecha korzystna (tekst dopasowuje się do możliwości ekranu), lecz przeważają aspekty negatywne, gdyż zmienność parametrów i geometrii sprawia, że trudno zapewnić wiarygodność udostępnianego dokumentu. Strony w języku HTML i jego odmianach już od

¹⁵ Online Community for the OpenDocument OASIS Standard - <http://opendocument.xml.org/> [2011.04.20].

¹⁶ HTML 4.01 Specification - <http://www.w3.org/TR/html4/> [2011.04.20].

¹⁷ World Wide Web Consortium (W3C)- <http://www.w3.org/> [2011.04.20].

¹⁸ Open Mobile Alliance - <http://www.openmobilealliance.org/> [2011.04.20].

początku XXI są wieku niemal wyłącznie pisane z zastosowaniem kodowania UTF-8, często używają języków skryptowych (np. JavaScript) i wykorzystują XML do reprezentacji swej struktury.

W odpowiedzi na powyższe słabości dokumentów tekstowych (w szczególności w zastosowaniu do e-administracji)¹⁹ upowszechniło się w ostatnich latach specyficzne podejście do tej kwestii. W nowoczesnych systemach teleinformatycznych dokumenty przechowuje się w **formacie strukturalnym** i na życzenie generuje odpowiedni format prezentacyjny. W takim systemie pierwotny dokument tekstowy zapisany jest w trzech plikach: dane (XML), struktura (XSD), wygląd (XSLT), na podstawie których na życzenie jest generowany plik finalny: PDF, HTML, RTF lub DOC. Zaletą tego rozwiązania jest duża otwartość na nowe standardy prezentacyjne, a nade wszystko oddzielenie danych od warstwy prezentacji, którą najczęściej jest urzędowy formularz.

4.2. Formaty graficzne

Grafikę w systemach informatycznych można reprezentować w formie wektorowej bądź rastrowej. W pierwszym modelu obiekty są tworzone z figur i krzywych matematycznych, w drugim – rysowane za pomocą punktów „pikseli” na płaszczyźnie w układzie kartezjańskim. Pierwszy rodzaj charakteryzuje się tym, że obrazy można dowolnie powiększać, bez utraty jakości; w drugim zaś granicę ową określa rozdzielczość, gdyż po przekroczeniu pewnej wartości granicznej powiększenia oraz staje się coraz mniej czytelny. Łatwo wydedukować, że każdy z tych sposobów do zakodowania informacji wymaga innych ilości danych. Zdecydowanie mniej w danych definiujących wymagają obrazy wektorowe, gdyż ilość parametrów niezbędnych do zdefiniowania figur geometrycznych jest niewielka (np. dla okręgu: współrzędne środka, promień, grubość linii i jej kolor). W modelu rastrowym natomiast zakodować należy współrzędne i atrybuty każdego piksela – stąd rozmiar pliku jest znaczny i rośnie proporcjonalnie do kwadratu rozdzielczości oraz wprost proporcjonalnie do innych parametrów (np. głębi koloru). Najistotniejsza jednak różnica pomiędzy grafiką wektorową a rastrową dotyczy warstwy prezentacji, a ściślej przenośności. Obrazy rastrowe dają się łatwo zapisać w formatach wymiennych i można je wyświetlać w rozmaitych programach, z kolei obrazy wektorowe funkcjonują poprawnie jedynie w aplikacjach

¹⁹ A. Matejko, J. Januszewski: E-paczka, czyli wysyłam dokumenty do urzędu - <http://www.podpiselektroniczny.pl/epaczka.htm> [2011.04.20]; Rozporządzenie Rady Ministrów z dnia 11 października 2005 r. w sprawie minimalnych wymagań dla systemów teleinformatycznych, „Dziennik Ustaw”. Nr 212, poz. 1766) – załącznik 2.

macierzystych, w których zostały stworzone (np. CorelDraw, AutoCad), czyli w formatach natywnych (odpowiednio: cdr, dwg). Cecha ta powoduje, że na potrzeby prezentacji wykorzystuje się wyłącznie grafikę rastrową, zaś finalne wersje obrazów wektorowych konwertuje do postaci rastrowej.

Głównymi parametrami opisującymi grafikę rastrową są: rozmiary obrazu (wysokość i szerokość liczona w pikselach); rozdzielczość (czyli liczba pikseli przypadająca jednostkę długości), liczona zwykle w pikselach na cal (ppi – pixels per inch) oraz parametr głębi koloru (color depth) określający paletę użytych kolorów, czyli liczbę bitów potrzebną do reprezentacji koloru danego piksela. W praktyce stosuje się cztery modele głębi: 1-bitową (czern i biel), 2-8-bitową (odcienie szarości [4-256 odcieni]), zaś kolor można reprezentować w palecie indeksowanej (2-8 bitów [4-256 kolorów]) lub w modelu RGB (6-24 bitów [64-16777216 kolorów]). Zagadnienia koloru obejmują więcej parametrów, które można analizować w ramach stosowanej przestrzeni barwnej. Dominującym sposobem reprezentacji koloru jest model RGB, czyli barwa addytywna. W przestrzeni tej każdy kolor uzyskiwany jest z mieszania trzech barw podstawowych (kanałów): R – red (czerwonej), G – green (zielonej) i B – blue (niebieskiej). Jeżeli głębię koloru dla każdego kanału wynosi 8 bitów, to pozwoli to na uzyskanie $2^{8+8+8}=2^{24}$ barw o 256 poziomach jasności, czyli ponad 16 milionów (true color). Powszechny jest także model CMYK, gdzie kolor uzyskiwany metodą subtraktywną, czyli nakładania barw: C – cyan (turkus), M – magenta (karmazyn), Y – yellow (żółty), K – Black (czarny). Obie przestrzenie są równie popularne, lecz mają odrębny zakres stosowania: RGB wykorzystuje się do prezentacji na monitorach, CMYK natomiast w działalności poligraficznej. Dodatkowym atutem tych rozwiązań jest możliwość bezstratnej konwersji barw RGB/CMYK, np. dla głębi 24-bitowej wg wzoru: $R=255-C$, $G=255-M$, $B=255-Y$ (K – nie podlega konwersji). Prócz przestrzeni RGB/CMYK barwę opisuje jeszcze uniwersalny model HSV (Hue, Saturation, Value), gdzie: H – hue (barwa światła), S – saturation (nasycenie koloru), V – Value (jasność, czasami określana jako B - brightness). Model ten nawiązuje do sposobu, w jakim barwę widzi ludzkie oko. H – to odcień, czyli cecha jakościowa barwy (np. zielony, niebieski, różowy), S – udział barwy czystej w kolorze (np. jasnoniebieski); V – intensywność barwy w zależności od wartości światła odbitego. Według tego modelu wszelkie barwy wywodzą się ze światła białego, gdzie część widma zostaje wchłonięta a część odbita od oświetlanych przedmiotów. Prócz wyliczonych uniwersalnych abstrakcji można jeszcze wskazać kilka innych parametrów o węższym zakresie stosowania, np. kanał alfa (czyli barwa przezroczysta), typowa dla palety

indeksowanej) oraz szereg opcji używanych do manipulacji na barwie, np. kontrast czy korekcja gamma.

Wśród licznych formatów stosowanych do prezentacji grafiki rastrowej (istnieje ich ok. 250)²⁰ szczególnie użyteczne w różnych zastosowaniach są cztery specyfikacje: JPEG, GIF, PNG i TIF. Najbardziej popularny jest pierwszy z nich. **JPEG** (JPG) jest formatem stratnym, który sprawdza się szczególnie dla obiektów złożonych z dużej liczby barw i dużej liczby szczegółów (np. pejzaże i portrety), charakteryzujących się płynnymi przejściami barw oraz brakiem lub małą ilością ostrych krawędzi i drobnych detali. Największą zaletą formatu jest mały rozmiar pliku wyjściowego przy zachowaniu akceptowalnej jakości oraz obsługa pełnej przestrzeni barwnej (true color); wadą zaś jednokierunkowa utrata jakości podczas kompresji. Algorytm kompresji JPG przekształca obraz w trzech etapach: najpierw każdy piksel (dla kolejnych kanałów RGB z osobna) zostaje przekształcony na skalę jasności (luminancję) oraz 2 barwy (chrominancje); następnie obraz jest dzielony na bloki 8x8 pikseli i poddany transformacji DTC (discrete cosine transform), która sprawia, że pikselom zostają przypisane uśrednione wartości barw (zmiennoprzecinkowe); na końcu następuje etap kwantyzacji, czyli transformacja wartości zmiennoprzecinkowych na całkowite (tylko tu powstają straty). Warto zauważyć, że skuteczność algorytmu zależy od zawartości obrazu, stąd niemal w każdym przypadku należy empirycznie dobrać optymalne ustawienia. W praktyce proces konwersji sprowadza się do doboru kompresji (0-100%) i obserwacji utraty jakości (granica akceptacji); jedynie w zastosowaniach profesjonalnych pomocne mogą być inne parametry: czynnik próbkowania i optymalizacja tabeli Huffmana – czyli parametry drugiego kroku kompresji oraz wygładzanie (antialiasing) i progresja (sukcesywne wyświetlanie obrazu w trakcie ładowania). Warto dodać, że JPG doskonale obsługuje popularne schematy metadanych (EXIF, IPTC i XMP), zaś ich wadą jest brak obsługi kanału alfa. Standard JPG ma wśród innych specyfikacji pozycję ustabilizowaną. Prace nad nim rozpoczęto stosunkowo dawno, bo już w roku 1983; najpierw pod egidą ISO, potem w ramach specjalnej grupy (Joint Photographic Experts Group)²¹. Pierwsza specyfikacja powstawała w latach 1991-1998 (ISO/IEC IS 10918-1), zaś ostatnia, w której zawarto szereg poprawek ogłoszono w 1999 (ISO/IEC IS 14495-1). Rok później na bazie dotychczasowych doświadczeń opracowany nowy standard **JPG2000** (norma ISO/IEC 15444-1), w którym wykorzystano nową metodę transformacji (DWT – kafelkową) oraz wbudowano możliwość

²⁰ Por. zestawienie: http://en.wikipedia.org/wiki/Comparison_of_graphics_file_formats [2011.04.20].

²¹ Joint Photographic Experts Group - <http://www.jpeg.org/> [2011.04.20], tamże dostępne wszystkie specyfikacje i raporty techniczne.

stosowania kompresji bezstratnej. Format ten nie zyskał jednak większej popularności, gdyż opcja kompresji bezstratnej tylko nieznacznie przewyższa popularny algorytm LZW stosowany np. w formacie TIFF.

Drugim popularnym formatem rastrowym jest **GIF** (Graphics Interchange Format), którego wykorzystanie daje najlepsze efekty w obrazach złożonych z małej liczby barw i zawierających duże przestrzenie jednokolorowe (np. obrazy konwertowane z grafiki wektorowej); sprawdza się także przy obiektach zawierających ostre krawędzie i drobne detale (np. czcionki). GIF uznawany jest za format stratny, co tylko częściowo jest prawdą. Stratność wynika z faktu, że obsługuje on tylko kolor 8-bitowy (256 barw) w jednym bloku (ramce). Teoretycznie możliwe jest więc podzielenie obrazu na bardzo małe bloki, co prowadziłoby do poszerzenia palety. Jednak większość programów graficznych stosuje jedynie jedną ramkę, co ostatecznie sprawia, że w praktyce GIF obsługuje maksymalnie 256 barw. Poza ograniczeniem palety GIF nie zniekształca obrazu, gdyż w obrębie bloku kompresja ma charakter bezstratny (LZW). Ciekawą jego własnością jest obsługa przezroczystości, czyli kanału alfa (transparent background color). Funkcjonalność ta sprawia, że w ramach formatu można tworzyć proste animacje (animowane GIFy) i obrazy z przezroczystym tłem (np. dla znaków wodnych). Druga cecha: przeplot (interlacing) – pozwala wyświetlać na grafikę warstwa po warstwie (4 fazy), w praktyce sprawia to, że na początku obraz wydaje się rozmyty, a w trakcie ładowania nabiera ostrości. GIF został stworzony w 1987 roku przez firmę CompuServe, a w 1989 doczekał się aktualizacji, dodającej między innymi możliwości przechowywania w nim animacji (Gif89a)²². Przy opracowywaniu specyfikacji, firma zaniedbała problem użytego algorytmu kompresji LZW, który był chroniony patentami firm IBM oraz UniSys. W połowie lat 90. nastąpił regres, gdyż UniSys zapowiedziała, że będzie ścigać firmy używające ich chronionej technologii i pobierać od nich opłaty licencyjne. Stan niepewności trwał aż do 2006 roku, gdy ostatecznie patenty wygasły. Równolegle w odpowiedzi na istniejące zagrożenie podjęto prace nad następcą GIFa – formatem **PNG**. Nowy standard oferuje funkcjonalność zbliżoną do GIF, choć na jego korzyść przemawia możliwość stosowania pełnej palety barwnej (true color). W praktyce jednak PNG nie zastąpił GIFa, ani nie zagroził JPG i TIF, gdyż oferowane opcje kompresji stratnej dają nieco gorsze efekty niż GIF/JPG, zaś algorytm bezstratny okazuje się tylko nieznacznie lepszy od TIF. PNG sprawdza się jednak na wielu specyficznych obszarach (np. zrzuty ekranowe lub obiekty o dużych rozmiarach), jest ponadto formatem domyślnym w

²² Graphics Interchange Format, Version 89a - <http://www.w3.org/Graphics/GIF/spec-gif89a.txt> [2011.04.20]

środowisku Linux i rekomendowanym przez W3C do zastosowań internetowych. Od 1995 roku nad rozwojem formatu czuwa Internet Engineering Steering Group (IESG), wspierana przez Konsorcjum W3C; od 2003 roku PNG został uregulowany normą ISO/IEC 15948²³.

W grupie formatów prezencyjnych o wąskim zakresie stosowania należy także umieścić **TIF (G4)**. Został on wprowadzie zaliczony do formatów archiwalnych (por. omówienie w rozdz. 3), niemniej jego funkcjonalność jest znacznie szersza. W jednym zastosowaniu TIF okazuje się bezkonkurencyjny: obrazy czarno-białe (1-bitowe) z bezstratną kompresją CCITT Group 4 (w skrócie G4). W tym zastosowaniu obrazy są ok. 6-, 15-krotnie mniejsze od oryginału i jednocześnie znacznie mniejsze niż w innych formatach (GIF, PNG, JPG). Funkcjonalność ta sprawia, że TIF (G4) doskonale sprawdza się jako kontener na skany bitonalne. Warto jednak zaznaczyć, że uzyskanie dobrej jakości obrazu 1-bitowego możliwe jest bezpośrednio tylko dla grafiki kreskowej lub nieciągłotonalnej. W przypadku obrazów ciągłotonalnych (gdy przejście między barwami odbywa się stopniowo) należy podczas konwersji głębi zastosować któryś z algorytmów rozpraszania (ditheringu), np. algorytm Floyda-Steinberga, wzór lub rozpraszane adaptacyjne. W efekcie obrazy ciągłotonalne zostaną zamienione na odpowiednio zagęszczone kropki lub linie, które będą imitować pierwotne barwy. Dodajmy, że technika ta jest podstawą przygotowania obrazów do druku, z wykorzystaniem tzw. technologii rastra. Obrazy z rozproszeniem cechuje efekt mory (moiré), którą ludzkie oko postrzega jako deformację przy pewnych wielkościach powiększania obrazu, tzn. gdy między pikselami dochodzi do interferencji. Warto też dodać, że do oglądania obrazów tego typu należy używać przeglądarek, które obsługują opcję antyaliasingu (tj. programowego rozmycia), gdyż wzór rastra składa się z małych ostrych elementów, które dla ludzkiego oka sprawiają wrażenie dyskomfortu i powodują, że obraz wydaje się błady i rozmyty.

4.3. Formaty hybrydowe

Formaty hybrydowe to grupa specyfikacji, która łączy cechy formatów tekstowych i graficznych. W praktyce można wskazać dwa takie rozwiązania: PDF i DJVu. Pierwsze w systematykach MIME należy do grypy aplikacyjnej, drugie do graficznej.

²³ Portable Network Graphics. Specification (Second Edition) - <http://www.w3.org/TR/2003/REC-PNG-20031110/> [2011.04.20].

PDF (Portable Document Format)²⁴ jest implementacją języka PostScript, który z czasem wzbogacono o nowe możliwości (elementy hipertekstowe i aplikacyjne). W najszerszym sensie PDF to specyfikacja, która w zależności od wersji może pełnić rozmaite funkcje: prezentacji, przenoszenia, drukowania i przechowywania treści tekstowo-graficznych. Dokumenty PDF cechują się pełną niezależnością od platformy (w każdym środowisku prezentują się identycznie) oraz pełną niezależnością od źródła (zawierają w pliku wszelkie komponenty niezbędne do prawidłowego działania, jak czcionki i obrazy); mają nadto wbudowany mechanizm programowego powiększenia (zoom). Twórcą formatu była firma Adobe, która w latach 1993-2007 roku samodzielnie rozwijała jego kolejne wersje (od 1.1 do 1.7 i odpowiadające mu wersje Acrobat: 1.0-9.0), które sukcesywnie zyskiwały nowe funkcjonalności, np. podpis elektroniczny (wersja 1.3 – 2000), opcje 3D (1.6 – 2004); w 2008 roku zaszła istotna zmiana, gdyż firma ogłosiła publiczne otwarcie formatu i przekazała jego specyfikację organizacji Association for Information and Image Management (AIIM), która opracowała odpowiednią normę ISO 32000-1:2008 i przejęła obowiązki koordynacyjne. Pierwotnie zastosowaniem PDF była poligrafia i rozwijany z tą myślą format ewoluował do wersji X (PDF/X), która od 2001 roku stałą się podstawą normy ISO 15929 i ISO 15930. Kolejnym polem zastosowania było kontenerowanie, tj. długotrwałe przechowywanie danych – wersja X (PDF/X - ISO 19005: 2005) oraz potrzeby inżynierskie – wersja E (PDF/E - ISO 24517: 2008). W 2010 powstała specyfikacja na potrzeby transakcyjne VT (PDF/VT - ISO 16612-2), czyli do zastosowań przy wymiana danych w formularzach PDF, np. w e-administracji. Wciąż w trakcie rozwoju jest natomiast najistotniejsza z punktu widzenia bibliotek cyfrowych – wersja PDF/UA (projekt ISO/DIS 14289-1)²⁵ – czyli PDF uniwersalny na potrzeby prezentacji (universal accessibility).

PDF jest formatem bardzo złożonym. W plikach może być zawarty zarówno tekst, jak i grafika (rastrowa, wektorowa), a także elementy aktywne. Specyfikacja potrafi bardzo precyzyjnie obsłużyć wszelkie związane z tym kwestie. Dokumenty PDF umożliwiają zatem pełną obsługę specyficznych takich problemów, jak: kodowanie znaków (w tym UTF-8), rodzaje czcionek (Type1, TrueType, OpenType), atrybuty graficzne (przestrzeń barwna, profile itp.), formularze, strukturę XML, metadane, podpis elektroniczny i in. Z punktu widzenia bibliotek cyfrowych warto zwrócić uwagę na trzy kwestie: PDFy powstałe na bazie dokumentów tekstowych (adnotowanych lub nieadnotowanych) zyskują na wiarygodności i

²⁴ Dokumentacja na witrynie Adobe: <http://www.adobe.com/pl/products/acrobat/standards.html> [2011.04.18] oraz PDF Standards: <http://pdf.editme.com/> [2011.04.18].

²⁵ PDF Universal Accessibility Committee: <http://pdf.editme.com/PDFUA> [2011.04.18]

stają się pełni przenośne, bowiem pokonany został problem labilności wyglądu strony (na każdej platformie prezentują się identycznie); zaś dokumenty graficzne zyskują wyższy poziom zorganizowania (strony, bookmarki). Specyfikacja PDF w odróżnieniu od formatów tekstowych i graficznych stwarza też spore możliwości ochrony dostępu do treści (szyfrowanie, hasła, certyfikaty i in.), które sprawiają, że dokumenty te doskonale nadają się do serwowania treści udostępnianej w sposób limitowany. Powstaje też – podkreślmy – możliwość prezentowania grafiki wektorowej, bez ich konwersji do form rastrowych. Jedynym problemem, jaki może się pojawić w wyżej wymienionych zastosowaniach to wzrost objętości PDF względem formy pierwotnej. Aby zredukować objętość plików należy podczas konwersji zastosować specjalne opcje (np. dowsampling) lub skorzystać ze specjalnych narzędzi umożliwiających optymalizację pliku (optimize). PDF ma ponadto wiele innych funkcjonalności (np. warstwę komentarzy), jednak ich rola w bibliotekach cyfrowych jest niewielka. Odrębnym wątkiem związanym z plikami PDF jest ich edycja. Jakkolwiek jest ona możliwa w odniesieniu do tekstu i grafiki (Adobe Acrobat, Adobe Bridge), stosuje się ją raczej rzadko, i niemal wyłącznie w zastosowaniach poligraficznych. Na potrzeby bibliotek cyfrowych najczęściej generuje się pliki (zwykle drukarkami wirtualnymi) i w razie potrzeby dokonuje ich montowania. Czasami też wykonuje się edycję w trybie ograniczonym: dodawanie warstwy OCR, bookmarki, kadrowanie, metadane, obracanie i reformatowanie stron itp. Warto dodać, że w niektórych kręgach zamiast PDF stosowany jest czysty **PostScript** (PS)²⁶. Jest on szczególnie popularny wśród matematyków i chemików, czyli w środowiskach, które używają niestandardowych znaków oraz dedykowanych edytorów. Od lat prym w tej roli dzierży TeX (LaTeX)²⁷ lub jego klony. Efektem pracy w tym środowisku są albo dokumenty w formacie natywnym (DVI), albo wyeksportowane PostScriptu. Warto na zakończenie dodać, że aplikacje do tworzenia PDF dzielą się na dwie duże grupy: komercyjne, oparte na rozwiązaniach Adobe oraz bezpłatne, bazujące na kompilatorach PostScriptu (np. GhostScript)²⁸.

Drugim standardem hybrydowym jest **DJVu** - technologia przeznaczona do kompresji stratnej dokumentów skanowanych. W odróżnieniu od PDF, który jest kontenerem, DJVu opiera się na metodzie segmentacji obrazu i wydzieleniu warstw; dodatkowo istnieje możliwość osadzenia niesformatowanego tekstu. W uproszczeniu proces kompresji DJVu polega na tym, że w zależności od typu skanu obraz, jest analizowany i dzielony na kilka

²⁶ PostScript language reference. Third edition. Reading : Addison-Wesley Publishing Company, 1999 - <http://partners.adobe.com/public/developer/en/ps/PLRM.pdf> [2011.04.19].

²⁷ LaTeX – A document preparation system - <http://www.latex-project.org/> [2011.04.19].

²⁸ Home Page for Ghostscript - <http://pages.cs.wisc.edu/~ghost/> [2011.04.19].

warstw tła o niższej rozdzielczości i warstwę treści o wyższej rozdzielczości. Podczas wyświetlania warstwy nakładają się na siebie i tworzą iluzję ostrego obrazu. Istotną cechą omawianego formatu jest bardzo wysoki współczynnik kompresji, znacznie przewyższający nawet inne algorytmy – co sprawia, że pliki DJVu osiągają rozmiary najmniejsze z możliwych (nawet 10-20 razy mniejsze od JPG czy GIF). W odróżnieniu od PDF używanie technologii DJVu jest znacznie trudniejsze, gdyż do uzyskania dobrego efektu wymaga określenia wielu parametrów (np. dotyczących segmentacji), a także biegłego operowania profilami i innymi zmiennymi. Prace nad technologią DJVu rozpoczęto w 1996 r. w koncernie AT&T.; po 2000 roku jej rozwój odbywa się równolegle w dwu środowiskach: komercyjnym (LizardTech Inc., od 2008 Celartem Technology Inc.) oraz społecznym na licencji (projekt DJVuLibre na licencji GNU GPL)²⁹, gdzie czołową rolę odgrywa były pracownik AT&T Léon Bottou. Dwutorowość rozwoju sprawia, że użytkownicy mają do wyboru w pełni funkcjonalne aplikacje zarówno komercyjne (np. DocumentExpress), jak i ich bezpłatne odpowiedniki (DJVuLibre) oraz szereg przeglądarek i narzędzi³⁰. Szerzej na temat formatu DJVu poświęcony jest odrębny rozdział w niniejszej monografii.

4.4. Formaty audio

Dźwięk jest w systemach informatycznych reprezentuje sygnał dyskretny, który jest odwzorowaniem fali akustycznej. Wskazana transformacja jest realizowana w specjalnych urządzeniach (przetwornikach analogowo-cyfrowych), które ciągły przekształcają sygnał analogowy w nieciągły sygnał cyfrowy; zatem głównym parametrem charakteryzującym jakość dźwięku cyfrowego jest częstotliwość próbkowania, inaczej (sampling). Parametr ten winien być co najmniej dwa razy wyższy od pasma słyszalnego. Ponieważ ludzkie ucho słyszy dźwięki w zakresie ok. 16 Hz do 18 kHz, próbkowanie winno wynosić co najmniej 36 kHz; w praktyce dla jakości CD stosuje się wartość nieco wyższą – 44,1 kHz. Drugim parametrem, który w jeszcze większym stopniu wpływa na jakość odtwarzanego dźwięku, jest przepływność (bitrate). Parametr ten wskazuje na liczbę bitów potrzebną do cyfrowej reprezentacji jednej sekundy nagrania i jest wyrażany w bitach na sekundę (b/s=bps) i ich krotnościach (kbps, Mbps); im wyższa jego wartość, tym lepsza jakość (np. dla MP3 przyjmuje się najczęściej 128 kbps, rzadziej 192 kbps). Przy zbyt agresywnej kompresji, co skutkuje zmniejszeniem wartości bitrate sygnał staje się zniekształcony przez tzw. artefakty, czyli dźwięki, których nie ma w pliku oryginalnym. Aby obejść ów niekorzystny efekt w

²⁹ DJVULibre - <http://djvu.sourceforge.net/> [2011.04.19].

³⁰ DJVU.org - <http://djvu.org/resources/> [2011.04.19].

większości formatów stosuje się zmienną bitrate (VBR – variable bit rate), co wpływa na zmniejszenie rozmiarów plików. W fali dźwiękowej (i jej reprezentacji cyfrowej) możemy wyróżnić jeszcze inne składniki widma dźwięku i ich jednostki, lecz nie ich związek z formatami zapisu jest niewielki; należą tu m.in.: wysokość (Hz), głośność (dB), składowe, wysokość i barwa; podobnie niewielki wpływ wywierają standardy i modele, które wpływają jedynie na liczbę równoległych ścieżek, np. mono (1), stereo (2), Dolby surround (4-5), Dolby Digital (5 + 1) i in. Z uwagi na znaczne rozmiary nieskompresowanego sygnału większość formatów stosuje rozmaite algorytmy kompresji (kodeki). W efekcie na opis formatu audio zawsze składa się kontener multimedialny (container) oraz stosowany kodek.

Wśród ok. 40 formatów do zapisu dźwięku³¹ najpopularniejszym formatem stratnym audio stosowanym w internecie jest **MP3** (MPEG-1 Audio Layer 3), który doskonale sprawdza się w każdym zastosowaniu. Standard ów charakteryzuje się przede wszystkim znakomitym współczynnikiem kompresji, który w porównaniu z WAV jest 12-, 13-krotnie mniejszy. Istota kodowania MPEG jest dość złożona, w dużym uproszczeniu polega ona dzieleniu sygnału na tzw. ramki (frame), w obrębie których dokonywana jest kompresja (maskowanie i usuwanie pewnych danych); ponadto kodowane są różnice między sąsiednimi ramkami. Sam mechanizm kompresji MPEG był stworzony z myślą o filmach, ale praktyka pokazała, że sprawdza się także doskonale w funkcji kodeka audio. Prace nad formatem rozpoczęto w 1991 roku; głównym promotorem działań była grupa Motion Picture Expert Group (MPEG)³², wspierana przez naukowców z Instytutu Fraunhofera z Erlangen³³. W krótkim czasie standard okrzepł i już w 1993 został utrwalony jako norma ISO/IEC 11172-3. Podejmowane później próby prac na następcą standardu (np. MP3Pro) nie powiodły się. Jedynym wyjątkiem, który ostatnio zanotował umiarkowane powodzenie był powstały w 1997 roku format **AAC** (Advanced Audio Coding), oferujący lepszą jakość dźwięku przy podobnym rozmiarze danych. Standard jest oparty na kompresji MPEG-4 Part 14 i opisany w normie ISO/IEC 13818-7. Znalazł on szerokie zastosowanie w urządzeniach mobilnych (np. telefonach komórkowych, iPhone, iPodzie) oraz konsolach. W wyniku sporów patentowych między Microsoftem a Instytutu Fraunhofera powstał trzeci popularny format **WMA** (Windows Media Audio), bardzo zbliżony pod względem funkcjonalności do MP3. Niezależnie testy udowadniają, że przy niskim bitrate (do 96 kbps) jakość dźwięku jest lepsza niż MP3, zaś powyżej tej granicy standard MP3 daje lepsze efekty.

³¹ Comparison of audio formats - http://en.wikipedia.org/wiki/Comparison_of_audio_formats [2011.04.18].

³² MPEG.org - <http://www.mpeg.org/> [2011.04.18].

³³ Fraunhofer-Institut für Sichere Informationstechnologie - <http://www.sit.fraunhofer.de/> [2011.04.18].

W odrębny sposób odtwarzane są tzw. media strumieniowe, czyli takie, które oferują dostęp w czasie rzeczywistym do medium ciągłego (analogicznie jak transmisja radiowa). Wyróżnia się przy tym dwa rodzaje dostępu: na żywo (ang. live), gdy transmitowane są dane, które od razu docierają do wszystkich użytkowników, np. transmisja na żywo lub na życzenie (ang. demand), gdzie serwer ma bazę zarejestrowanych danych, zaś użytkownik, wybiera sobie, który pokaz multimedialny od jakiego momentu chce oglądać. Do obsługi takich funkcjonalności stworzono specjalne formaty. Najpopularniejszym jest opracowany w 1995 roku przez firmę RealNetworks specyfikacja **RealAudio (RA, RM)**, który doskonale sprawdza się na słabych łączach internetowych. Wśród innych tego rodzaju warto wspomnieć o uniwersalnym formacie audio/wideo **QuickTime** rozwijaną przez Apple czy opracowanym przez Xiph.org kontenerze strumieniowym **Ogg**, który pozwala na wykorzystanie rozmaitych kodeków (większą popularnością cieszą się: uniwersalny kodek Vorbis oraz dedykowany dla mowy ogólnego Speex).

Trzecią grupą są formaty służące do rejestracji i odtwarzania dźwięków tworzonych sztucznie w środowisku komputerowym (w tym sztucznej mowy). Są one – mówiąc metaforycznie – dźwiękowym odpowiednikiem kolorów indeksowanych. Liderem wśród formatów tego typu jest stworzony w 1983 roku format **MIDI** skrót od (Musical Instrument Digital Interface)³⁴. Pliki tego rodzaju są bardzo małe i prócz kodowania dźwięku (próbki) zawierają komendy sterujące, które pozwalają na współpracę z zewnętrznymi urządzeniami muzycznymi.

4.5. Formaty wideo

Cyfrowa reprezentacja wideo (ruchomych obrazów) stanowi połączenie technologii reprezentacją grafiki rastrowej z problemami transmisji oraz technologią audio, która może być kodowana wspólnie lub na niezależnych ścieżkach. Wśród wielu pojęć i parametrów charakterystycznych dla tego obszaru można wskazać: przepływność (bitrate), pełniącą analogiczną funkcję jak dla dźwięku [por. 4.4], rozdzielczość i problemy związane z kolorem – jak dla grafiki rastrowej [4.2], lecz przede wszystkim parametr kluczowy dla filmów – czyli przepływność (framerate, frame frequency), która jest miarą prędkości (płynności) wyświetlania ruchomych obrazów wyrażoną w klatkach na sekundę (FPS – frames per second). Wartość przepływności w telewizji i nagraniach filmowych jest stała i zależy od sposobu kodowania. Zbyt niska wartość FPS powoduje u widza wrażenie „skakania” obrazu,

³⁴ MIDI Manufacturers Association - <http://www.midi.org/> [2011.04.18].

z kolei wysoki zwiększa wrażenie płynności wyświetlania. Z badań nad postrzeganiem wynika, że w przypadku ludzkiego oka w zupełności wystarczy wartość 24 FPS (urządzenia HD), 25 FPS (odtwarzacze DVD) lub 30 FPS (urządzenia mobilne). Warto jeszcze doprecyzować rozdzielczość (liczoną w liniach, która w przypadku wideo określają poszczególne standardy: VGA (640x480), PAL (768x576), XGA (1024x768) oraz HDTV w dwu wariantach (HDRedy 1280x720) lub (FullHD 1920x1080). Wynikają stąd znaczne różnice zapotrzebowania na dane, proste obliczenia pozwalają stwierdzić, że najmniejszy monitor VGA wyświetla jednorazowo 0,3 Mpix, zaś największy FullHD – prawie 2 Mpix. Nawet jeśli większość jest nadawana z przeplotem (na zmianę wyświetlane są linie parzyste i nieparzyste), co redukuje ilość danych o połowę – to i tak wideo w porównaniu z innymi mediami generuje gigantyczne zapotrzebowanie na transfer. W praktyce istnieją dwa wyjścia, aby ograniczyć bitrate przy zachowaniu akceptowalnej jakości: zastosować agresywną kompresję stratną (kodeki) oraz wykorzystać formaty wspierające skalowalność [ekranu]. Warto dodać, że formaty wideo, podobnie jak audio składają się zazwyczaj z kontenera multimedialny (istnieje ich ok. 25, np. MPEG-4 Part 14, AVI, ASF, 3GP, QuickTime, VOB, Ogg)³⁵, w którym zastosowany jest odpowiedni kodek, osobny dla obrazu (powstało ich ok. 35, np. H.264, MPEG-2, MPEG-4 Part 2, XVID)³⁶ i osobny dla dźwięku (np. MP3, AC3, AAC).

Na obecnym poziomie rozwoju technologii najpopularniejszy jest grupa formatów z rodziny MPEG, na której rozwojem czuwa Moving Picture Experts Group³⁷ - grupa robocza ISO/IEC zajmująca się rozwojem standardów kodowania audio i wideo, a także zatwierdzona przez ISO grupa powszechnie stosowanych formatów zapisu danych zawierających obraz i dźwięk. Aktualnie w praktyce telewizji cyfrowej używa się dwu formatów z tej grupy: MPEG-2, który opracowano w 1994 (norma ISO/IEC 13818-2:1996) oraz MPEG-4, który rozwija się od 1998 roku (norma ISO/IEC 14496 – 27 arkuszy). Specyfikacja **MPEG-2** obsługuje wszelkie rozdzielczości (352x288 do 1920x1152), pozwala na kodowanie obrazu zarówno z przeplotem, jak i bez niego, obsługuje skalowanie i jest zoptymalizowana dla przepływności 4 Mbps. Jakkolwiek jest on wciąż bardzo popularny z wolną jego traci użytkowników na rzecz standardu **MPG-4 [AVC]** (dokładnie MPG-4 Part 10 z kodekiem H.264), zwanego też AVC (Advanced Video Coding). Głównymi zaletami AVC jest koncepcja obiektowa, zorientowanie na media strumieniowe, skalowalność (176x144 do

³⁵ Comparison of container form. - http://en.wikipedia.org/wiki/Comparison_of_container_formats [2011.04.18].

³⁶ Comparison of video codecs - http://en.wikipedia.org/wiki/Comparison_of_video_codecs [2011.04.18].

³⁷ Moving Picture Experts Group - <http://mpeg.chiariglione.org/> [2011.04.18], także dokumentacja.

4096x2304), a nade wszystko doskonały współczynnik kompresji dzięki użyciu wielu innowacji oraz niezależność od medium i interakcja z użytkownikiem. Aktualnie istnieją dwaj główni konkurenci AVC: format **WMV** opracowany przez Microsoft, przyjęty przez amerykańskie stowarzyszenie SMPTE pod nazwą VC-1, oraz **AVS** opracowany przez konsorcjum stworzone przez rząd chiński. Wśród wielu innych formatów o funkcjonalności zbliżonej do H.264/AVC warto wspomnieć o popularnych kontenerach: **MOV** (znanej jako QuickTime firmy Apple), wspomnianym już wcześniej formacie **Ogg** oraz specyfikacji **FLV** (Flash Video), rozwijanej przez Adobe. Ostatnia specyfikacja stała się szczególnie popularna za sprawą używania jej w globalnych serwisach wideo: YouTube, Google Video, Reuters.com, Yahoo! Video czy MySpace.

4.6. Inne formaty

Wśród znacznej grupy innych formatów warto osobno podkreślić rozwijany przez Adobe format **SWF** (Shockwave Flash Object)³⁸, który są dominującą formą prezentacji w sieci animacji i interaktywnych apletów. Funkcjonalność i rozbudowę animacji fleszowych umożliwia ponadto udostępniony w 2004 roku język programowania **ActionScript**³⁹, który pozwala programować zdarzenia. Dzięki właściwościom języków skryptowych można łatwo tworzyć na stronach www rozmaite formy prezentacji dowolnego typu danych, np. grafiki rastrowej, dla której zaprogramowane są mechanizmy nawigacji, powiększania i efekty przejścia⁴⁰. Niektóre rodzaje danych ze względu na posiadane parametry lub potrzeby użytkowników posługują się dedykowanymi formatami. Przykładem jest serwowanie map – czyli obiektów o znacznych rozmiarach. Popularnym rozwiązaniem w tym zakresie okazało się podzielenie grafiki rastrowej na wiele części (nawet tysięcy) i ich ponowne wirtualne składanie w zależności od wybranej skali powiększenia. Na takim mechanizmie opiera się m.in. serwis GoogleMaps oraz popularny format **Zoomify**⁴¹. Innym przykładem są nuty. Dedykowany format **Sibelius**⁴², który odtwarza jednocześnie zapis i podkład muzyczny.

Wiele dedykowanych formatów stworzono także z myślą o urządzeniach mobilnych. Jednak jedynie kilka zdobyło ugruntowaną pozycję i jest prezentowanych w sieci dystrybucji e-booków, np. Powel's Books. Stosuje się albo formaty natywne: Microsoft Reader (**LIT**); MobiPocket (**PRC**), iPad (**iTunes**), <http://www.ebooks.com/help/reader->

³⁸ SWF File format specification (v. 10) - pdf/swf_file_format_spec_v10.pdf [2011.04.21].

³⁹ AAS - <http://www.adobe.com/devnet/actionscript/references.html> [2011.04.21].

⁴⁰ Przykład: <http://www.paulvanroekel.nl/picasa/flashgrid/demo/index.html> [2011.04.21].

⁴¹ Zoomify - <http://www.zoomify.com/> [2011.04.20].

⁴² Sibelius - <http://www.sibelius.com/> [2011.04.20].

software/mobipocket.asp Adobe Digital Editions (**PDF/DRM**) albo formaty otwarte: **EPUB**, **PDF** lub **Flash**. Ostatnia forma jest szczególnie interesująca, gdyż pozwala serwować dowolny typ dokumentu (np. PDF, TXT) i precyzyjnie zarządzać DRM. Rozwiązanie takie stosuje np. polska księgarnia Virtuallo oraz niektóre portale społecznościowe, np. Sciribd⁴³ czy DocStock⁴⁴.

Bibliografia

- Jaworski, R. (2009): *Multimedia i grafika komputerowa*. Warszawa: Wydawnictwa Szkolne i Pedagogiczne
- Kamiński, B. (2005): *Przetwarzanie informacji. Cz. 1, Przetwarzanie tekstu, przetwarzanie obrazu*. Warszawa: Wydawnictwa Szkolne i Pedagogiczne
- Owczarz-Dadan, A. (2006): *Tworzenie filmów w Windows XP*. Gliwice: Wydawnictwo Helion.
- Owczarz-Dadan, A. (2010): *Photoshop CS4 PL*. Gliwice: Wydawnictwo Helion
- Rudny, T. (2010): *Multimedia i grafika komputerowa*. Gliwice: Wydawnictwo Helion.
- Tamborski, T. (2004): *Przetwarzanie informacji. Cz. 2, Przetwarzanie dźwięku, przetwarzanie strumieniowe*. Warszawa: Wydawnictwa Szkolne i Pedagogiczne
- Ulrich, K. (2008): *Flash CS3 Professional PL*. Gliwice: Wydawnictwo Helion.
- Zakrzewski, P. (2007): *Adobe Premiere Elements*. Gliwice: Wydawnictwo Helion.
- Zimek, R. (2005): *ABC grafiki komputerowej*. Gliwice: Wydawnictwo Helion.

⁴³ Sciribd - <http://www.scribd.com/>

⁴⁴ Docstoc - <http://www.docstoc.com/>